RESULT 1
A38084
galactose oxidase (EC 1.1.3.9) precursor [validated] - fungus (Cladobotryum dendroides)
C;Species: Cladobotryum dendroides
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C;Accession: A38084
R;McPherson, M.J.; Ogel, Z.B.; Stevens, C.; Yadav, K.D.S.; Keen, J.N.; Knowles, P.F.
J. Biol. Chem. 267, 8146-8152, 1992
A;Title: Galactose oxidase of Dactylium dendroides. Gene cloning and sequence analysis.
A;Reference number: A38084; MUID:92235025; PMID:1569070
A;Accession: A38084
A;Molecule type: DNA
A;Residues: 1-728 <MCP>
A;Cross-references: GB:M86819
A;Note: it is uncertain whether Met-1 or Met-49 is the initiator
A;Note: parts of this sequence, including the amino end of the mature protein, were confirmed by protein sequencing
R;Ito, N.; Phililips, S.E.V.; Stevens, C.; Ogel, Z.B.; McPherson, M.J.; Keen, J.N.; Yadav, K.D.S.; Knowles, P.F.
Nature 350, 87-90, 1991
A;Title: Novel thioether bond revealed by a 1.7 angstrom crystal structure of galactose oxidase.
A;Reference number: A48244; MUID:91163641; PMID:2002850
A;Contents: annotation
R;Ito, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A;Reference number: A51740; PDB:1GOF
A;Contents: annotation; X-ray crystallography, 1.7 angstroms, residues 90-728
R;Ito, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A;Reference number: A51741; PDB:1GOG
A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 90-728
R;Ito, N.; Phillips, S.E.V.; Knowles, P.F.
submitted to the Brookhaven Protein Data Bank, September 1993
A;Reference number: A51742; PDB:1GOH
A;Contents: annotation; X-ray crystallography, 2.2 angstroms, residues 90-728
C;Genetics:
A;Gene: gaoA
C;Superfamily: Cladobotryum dendroides galactose oxidase
C;Keywords: disulfide bond; metal binding; oxidoreductase
F;1-64/Domain: signal sequence #status predicted <SIG>
F;65-89/Domain: propeptide #status predicted <PRO>
F;90-728/Product: galactose oxidase #status experimental <MAT>
F;107-116,604-607/Disulfide bonds: #status experimental
F;317-361/Cross-link: cysteinyltyrosine (Cys-Tyr) #status experimental
F;361,584,585,670/Binding site: copper (Tyr, Tyr, His, His) #status experimental
F;419/Binding site: substrate (Arg) #status predicted

```
Query Match              99.9%;   Score 3420;  DB 1;  Length 728;
Best Local Similarity    99.8%;   Pred. No. 5.7e-206;
Matches  638;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ASAPIGSAISRNNWAVTCDSAQSGNECNKAIDGNKDTFWHTFYGANGDPKPPHTYTIDMK 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     90 ASAPIGSAISRNNWAVTCDSAQSGNECNKAIDGNKDTFWHTFYGANGDPKPPHTYTIDMK 149

Qy     61 TTQNVNGLSMLPRQDGNQNGWIGRHEVYLSSDGTNWGSPVASGSWFADSTTKYSNFETRP 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    150 TTQNVNGLSMLPRQDGNQNGWIGRHEVYLSSDGTNWGSPVASGSWFADSTTKYSNFETRP 209

Qy    121 ARYVRLVAITEANGQPWTSIAEINVFQASSYTAPQPGLGRWGPTIDLPIVPAAAAIEPTS 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    210 ARYVRLVAITEANGQPWTSIAEINVFQASSYTAPQPGLGRWGPTIDLPIVPAAAAIEPTS 269

Qy    181 GRVLMWSSYRNDAFGGSPGGITLTSSWDPSTGIVSDRTVTVTKHDMFCPGISMDGNGQIV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    270 GRVLMWSSYRNDAFGGSPGGITLTSSWDPSTGIVSDRTVTVTKHDMFCPGISMDGNGQIV 329

Qy    241 VTGGNDAKKTSLYDSSSDSWIPGPDMQVARGYQSSATMSDGRVFTIGGSWSGGVFEKNGE 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    330 VTGGNDAKKTSLYDSSSDSWIPGPDMQVARGYQSSATMSDGRVFTIGGSWSGGVFEKNGE 389

Qy    301 VYSPSSKTWTSLPNAKVNPMLTADKQGLYRSDNHAWLFGWKKGSVFQAGPSTAMNWYYTS 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 VYSPSSKTWTSLPNAKVNPMLTADKQGLYRSDNHAWLFGWKKGSVFQAGPSTAMNWYYTS 449

Qy    361 GSGDVKSAGKRQSNRGVAPDAMCGNAVMYDAVKGKILTFGGSPDYQDSDATTNAHIITLG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    450 GSGDVKSAGKRQSNRGVAPDAMCGNAVMYDAVKGKILTFGGSPDYQDSDATTNAHIITLG 509

Qy    421 EPGTSPNTVFASNGLYFARTFHTSVVLPDGSTFITGGQRRGIPFEDSTPVFTPEIYVPEQ 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    510 EPGTSPNTVFASNGLYFARTFHTSVVLPDGSTFITGGQRRGIPFEDSTPVFTPEIYVPEQ 569

Qy    481 DTFYKQNPNSIVRVYHSISLLLPDGRVFNGGGGLCGDCTTNHFDAQIFTPNYLYNSDGNL 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db    570 DTFYKQNPNSIVRVYHSISLLLPDGRVFNGGGGLCGDCTTNHFDAQIFTPNYLYNSNGNL 629

Qy    541 ATRPKITRTSTQSVKVGGRITISTDSSISKASLIRYGTATHTVNTDQRRIPLTLTNNGGN 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    630 ATRPKITRTSTQSVKVGGRITISTDSSISKASLIRYGTATHTVNTDQRRIPLTLTNNGGN 689

Qy    601 SYSFQVPSDSGVALPGYWMLFVMNSAGVPSVASTIRVTQ 639
          |||||||||||||||||||||||||||||||||||||||
Db    690 SYSFQVPSDSGVALPGYWMLFVMNSAGVPSVASTIRVTQ 728
```

```
Database :        Issued_Patents_AA:*
                  1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                  2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                  3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                  4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                  5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                  6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*


     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

 # same mutations (handwritten)

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3420	99.9	639	4	US-09-782-906-2	Sequence 2, Appli
2	3401	99.3	639	4	US-09-782-906-3	Sequence 3, Appli
3	3391	99.0	639	4	US-09-782-906-4	Sequence 4, Appli
4	3390	99.0	639	4	US-09-782-906-5	Sequence 5, Appli
5	2207	64.4	679	3	US-09-257-536-2	Sequence 2, Appli
6	2207	64.4	679	3	US-09-512-230-2	Sequence 2, Appli
7	194	5.7	2736	4	US-09-252-991A-30227	Sequence 30227, A
8	170.5	5.0	1216	4	US-09-134-000C-5130	Sequence 5130, Ap
9	161	4.7	1752	4	US-09-865-621A-2	Sequence 2, Appli
10	150	4.4	1146	4	US-09-198-452A-580	Sequence 580, App
11	150	4.4	1749	4	US-09-640-419C-28	Sequence 28, Appl
12	143.5	4.2	3892	4	US-09-328-352-5503	Sequence 5503, Ap
13	137.5	4.0	642	4	US-09-337-307A-3	Sequence 3, Appli
14	137.5	4.0	642	4	US-09-337-307A-4	Sequence 4, Appli

(handwritten bracket grouping results 1–4 with notation "6498026")

```
Database :        SPTREMBL_25:*
                  1:  sp_archea:*
                  2:  sp_bacteria:*
                  3:  sp_fungi:*
                  4:  sp_human:*
                  5:  sp_invertebrate:*
                  6:  sp_mammal:*
                  7:  sp_mhc:*
                  8:  sp_organelle:*
                  9:  sp_phage:*
                 10:  sp_plant:*
                 11:  sp_rodent:*
                 12:  sp_virus:*
                 13:  sp_vertebrate:*
                 14:  sp_unclassified:*
                 15:  sp_rvirus:*
                 16:  sp_bacteriap:*
                 17:  sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1730	50.5	716	3	Q870R5	Q870r5 neurospora
2	452.5	13.2	779	2	Q938A2	Q938a2 stigmatella
3	416.5	12.2	526	2	Q53679	Q53679 stigmatella
4	414.5	12.1	757	16	Q82I66	Q82i66 streptomyce
5	337.5	9.9	645	16	Q9RDB4	Q9rdb4 streptomyce
6	300.5	8.8	656	16	Q81DE5	Q81de5 bacillus ce
7	299.5	8.7	645	16	Q82CX0	Q82cx0 streptomyce
8	298.5	8.7	615	10	Q9FYG4	Q9fyg4 arabidopsis
9	295	8.6	594	10	Q8RY19	Q8ry19 arabidopsis

Database :   SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	3420	99.9	680	1	GAOA_DACDE	Q01745	dactylium d
2	233	6.8	647	1	NANH_MICVI	Q02834	micromonosp
3	209	6.1	1014	1	NANH_CLOSE	P29767	clostridium
4	164	4.8	1723	1	KA93_HUMAN	Q9upq9	homo sapien
5	162	4.7	725	1	YA33_SULSO	Q97z97	sulfolobus
6	149	4.4	1723	1	PM20_CHLPN	Q9z812	chlamydia p
7	142.5	4.2	1902	1	P2P_LACPA	Q02470	lactobacill
8	141.5	4.1	751	1	KHL1_MOUSE	Q9ji74	mus musculu
9	141.5	4.1	959	1	N100_YEAST	Q02629	saccharomyc
10	140.5	4.1	827	1	XANP_XANS2	Q60106	xanthomonas
11	138.5	4.0	748	1	KHL1_HUMAN	Q9nr64	homo sapien
12	137	4.0	1356	1	HET1_PODAN	Q00808	podospora a
13	137	4.0	1656	1	OMPB_RICJA	O06653	r outer mem
14	136	4.0	1286	1	AIDA_ECOLI	Q03155	escherichia
15	135.5	4.0	916	1	SCRB_LIMPO	Q25386	limulus pol
16	135	3.9	2432	1	Y43R_IRV6	P18305	chilo iride
17	132	3.9	1147	1	TEA1_SCHPO	P87061	schizosacch
18	132	3.9	2003	1	YDBA_ECOLI	P33666	escherichia
19	131.5	3.8	1902	1	P2P_LACLC	P15293	lactococcus